United Nations E/CN.3/1999/4/Add.1

Economic and Social Council

Distr.: General
29 October 1998

Original: English

Statistical Commission
Thirtieth session
1–5 March 1999
Item 3 (b) of the provisional agenda*
International trade statistics

International trade statistics

Note by the Secretary-General

Addendum

Summary

The present note considers the possibilities for rationalization of the Commodity Trade Statistics Database (Comtrade), which contains data in terms of five separate commodity classifications. Basically, the aim is to determine whether a smaller set of data could be maintained, consistent with meeting user needs, by holding a smaller number of classifications in the database. It also presents ideas on the possible use of the Central Product Classification (goods part) for international merchandise trade statistics by the United Nations Statistics Division.

The note was prepared by the United Nations Statistics Division subsequent to the meeting of the Task Force on International Trade Statistics on 25 and 26 February 1998, and follows up the discussion and conclusions of the Task Force at that meeting (see E/CN.3/1999/4). It reflects comments made by all members of the Task Force on a draft that was circulated in October 1998.

The note is submitted to the Commission in response to a request of the Statistical Commission's Working Group on International Statistical Programmes and Coordination (see E/CN.3/1999/20, para. 10 (b)). It will be discussed by the Task Force at its next meeting, scheduled for 8 to 10 March 1999.

Points for discussion are set out in paragraph 17.

* E/CN.3/1999/1.

98-33273 (E) 111298

Contents

		Paragraphs	Page
I.	Introduction	1–3	3
II.	Views of the Task Force on International Trade Statistics	4–11	3
III.	Analysis	12–15	4
IV.	Other issues	16	5
V.	Points for discussion	17	5

I. Introduction

1. The Commodity Trade Statistics Database (Comtrade) is maintained by the United Nations Statistics Division, and contains detailed commodity by partner trade data for approximately 100 countries for each year from 1962 to date. These are held according to five commodity classifications: the Harmonized Commodity Description and Coding System (HS) in its 1996 revision (HS 96) (data from 1996), and in its 1988 revision (HS 88) (data from 1988), and the Standard International Trade Classification (SITC), Revision 3 (data from 1988), Revision 2 (data from 1976) and Revision 1 (data from 1962).

2. The original impetus towards rationalization of the Database arose from significant increases in the requirements for and consequent cost of disk space for storing data on-line, and simultaneous reductions of staff available to carry out the processing and conversion of data into the various classifications, and for maintenance of the Database. The Task Force had endorsed the general idea of rationalizing the Database under those circumstances, and some rationalization was achieved by ceasing the processing and storage of quarterly data, with considerable reduction in the reporting burden on countries. The Statistical Commission, in 1997, cautioned the United Nations Statistics Division to assess as fully as possible the needs of users of Comtrade before rationalizing its contents.[1] In 1998, the Statistical Commission's Working Group on International Statistical Programmes and Coordination endorsed the idea of rationalization, specifically in respect of the situation where data are maintained in terms of the five separate classifications, and requested the Task Force to study it further (see E/CN.3/1999/20, para. 10).

3. The Task Force attempted to develop a rationalization strategy taking account of user needs; this is summarized in the above-mentioned report of the Task Force (see E/CN.3/1999/4, para. 9), and is set out in greater detail below.

II. Views of the Task Force on International Trade Statistics

4. Members of the Task Force generally agreed that rationalization was possible. There was support from all organizations for the idea that irrespective of any rationalization, the Database should retain the data in the classification version originally reported by each country so that users could resort to that reported data for validation and other purposes (in cases where countries report in terms of a national classification that is not accommodated in Comtrade, the United Nations Statistics Division converts the reported data to the nearest HS or SITC version for storage in the Database). In all scenarios, there was agreement to remove data according to SITC, Rev. 3, the rationale being that HS data provide more detail and are convertible by users into SITC, Rev. 3 and other versions of SITC; in addition, the SITC, Rev. 3 data, being available from 1988, are not available for a sufficiently long period for some types of long-term analysis. However, there are known users of SITC, Rev. 3 data who could then be faced with the need to create the SITC, Rev. 3 data for themselves from HS data rather than being able to obtain it directly from Comtrade.

5. There is a characteristic of country reporting practices that affects Comtrade and makes it difficult to be firm about retaining only a given classification from a specific date. The characteristic is that not all countries adopt a classification in its initial year; some countries report the data to the United Nations Statistics Division according to a superseded and less detailed classification for one or more years after the recommended date of introduction of a new classification. For example, it is not feasible to have all HS 88 data commencing 1988, and to drop SITC, Rev. 3 from that year. If this were done, there would be gaps in the Database since many countries reported SITC, Rev. 3 data and not HS 88 data in 1988 and even later years. In addition, reverse conversion from SITC, Rev. 3 to HS is not sufficiently reliable to fill those gaps since one SITC, Rev. 3 code may be equivalent to the sum of several HS codes. This is a large-scale problem since in Comtrade there are 232 country periods for approximately 70 countries reported in SITC, Rev. 3 since 1988, the year of introduction of HS 88.

6. Views diverged as to the advisability of various options for rationalization. For long-term analysis, some organizations were of the opinion that SITC, Rev. 1 data should be maintained for all years (1962 to the current year) and SITC, Rev. 2 data should be removed from Comtrade. The World Trade Organization has provided an analysis in support of this option. This option provides good quality SITC, Rev. 1 data for the period 1962 to the current year since the conversion into SITC, Rev. 1 of data reported in SITC, Rev. 2 is good; however the needs of users of SITC, Rev. 2 data would not be met.

7. Other organizations considered that SITC, Rev. 2 data should be kept for long-term series, SITC, Rev. 1 data should be removed from Comtrade for the years where SITC, Rev. 2 data was available, and SITC, Rev. 1 data should be converted to SITC, Rev. 2 format where SITC, Rev. 2 was

not available. This option provides good quality SITC, Rev. 2 data from about 1976 to the current year since SITC, Rev. 2 data was introduced in about 1976. Prior to 1976, data were reported in SITC, Rev. 1; conversion of data reported in SITC, Rev. 1 into SITC, Rev. 2 has significant deficiencies (based on a study carried out by the United Nations Statistics Division in 1998), so that long-term time series in SITC, Rev. 2 would be fully reliable only from about 1976. The needs of users of SITC, Rev. 1 data would not be met under this option (however, as a refinement to this option, in an effort to meet needs for SITC, Rev. 1 data the original reported SITC, Rev. 1 data could be maintained in Comtrade from 1962 to 1975, as suggested in paragraph 4 above, and a conversion table of SITC, Rev. 2 to SITC, Rev. 1 data could be made available for users to derive their own SITC, Rev. 1 data from the SITC, Rev. 2 data in Comtrade for years 1976 to date).

8. For current and short-term analysis, a number of organizations considered that only HS 96 was needed, and that HS 88 and SITC, Rev. 3 could be removed from Comtrade (either SITC, Rev. 1 or SITC, Rev. 2 data would be retained for long-term analysis). Others considered that data reported in HS 96 could be removed and the HS 88 data converted from the HS 96 data could be retained, giving HS 88 data available from about 1988 to the current year. The former provides the most detailed information for recent years but only for a short period (1996 to the current year). The latter provides a time series since 1988 but does not present data at all in terms of the most current, more detailed HS classification, namely, HS 96; it is also inconsistent with the ideas expressed in paragraph 4 above concerning retention of data in the classification in which it was reported.

9. Other options for rationalization also exist, such as (a) processing and storing only the original classification in which the data were reported to the United Nations Statistics Division and converting that data into other needed classifications through computerized correlation tables at the time of each extraction of the data, or (b) processing and storing data according to all classifications but eliminating double entries at storing time.

10. The first option is not very practical. It would only work if users of the Database requested small numbers of commodities at a time. However, many users need to be able to extract large amounts of data at one time, which would take a lot of (expensive) CPU time for each extraction. Response times to meet requests would also be relatively slow.

11. The second option would leave the current situation intact from a user's point of view, i.e., data available according to five classifications, but would reduce the occupied disk space by about half. By storing all classifications individually, as is done currently, for certain commodities exactly the same trade information is stored up to four or five times in the Database due to the fact that particular commodity codes in one classification are equivalent to identical codes in each of the other classifications. For instance, "Beer made from malt" is represented by just one code in each of the five classifications maintained in Comtrade, namely 1123 in SITC classifications and 220300 in HS classifications. Maximum reduction of disk space would be achieved if duplication of the aggregated levels of the classification was also eliminated at storing time. If the software for compressing and decompressing the Comtrade data could be amended to achieve this, the storage space could be more effectively used without having to eliminate one or more classifications. Initial indications are that this would require a major re-engineering of the Database and extraction software.

III. Analysis

12. It is clear that whatever rationalization strategy were to be adopted, specific requirements of some current users of Comtrade would be frustrated and/or a considerable amount of new computer systems development work would have to be done. Although the automatic recording of usage by on-line users has not been modified to provide details of usage of the various classifications, a review of requests received from users and information from international organizations on their on-line usage indicates that among the current users of data from Comtrade, Governments and international organizations are users of HS 96, HS 88, SITC, Rev. 3, SITC, Rev. 2 and SITC, Rev. 1; universities and other research organizations are users of HS 88, SITC, Rev. 3, SITC, Rev. 2; and the private sector (ad hoc requests) are mainly users of HS 96, HS 88 and SITC, Rev. 3, with occasional use of SITC, Rev. 2 and Rev. 1 data for longer-term analysis. There is significant use at all levels of detail of the classifications as well, although indications are that SITC, Rev. 2 and Rev. 1 data tend to be used more at the more aggregated levels of those classifications and HS 96, HS 88 and SITC, Rev. 3 data at the more detailed levels. Some organizations indicated that they thought it would be useful to have trade data according to the Central Product Classification (CPC) and the International Standard Industrial Classification of All Goods and Services (ISIC).

Given correlation tables from HS or SITC to other classifications, such as CPC or ISIC, some users indicated that they were prepared to derive any such needed data.

13. Since the time that the idea of rationalization was raised, changes in the arrangements for computing services at the United Nations have led to a reduction in both disk space usage and its cost, and the processing load has been reduced by the elimination of quarterly data from the Database. Consequently, the situation is no longer acute, so that the need to rationalize is less urgent.

14. The computer system for Comtrade is quite robust, and is designed to accommodate data according to new revisions of classifications while retaining data according to older versions of these classifications. In principle, it could also accommodate data according to additional classifications, such as CPC (goods part) and ISIC. This would involve some development cost (both for computer systems and substantive work), staff resources to provide reliable conversions from HS or SITC data to the new classifications, and operating costs to process and convert data and maintain the Database.

15. Members of the Task Force are further reviewing user needs and taking other steps discussed in the report of the Task Force (see E/CN.3/1999/4, para. 9), and the results will be reviewed at the Task Force meeting scheduled for 8 to 10 March 1999.

IV. Other issues

16. In response to other elements of the request of the Working Group (see E/CN.3/1999/20, para. 10 (b)) the following information is presented:

(a) The international organizations normally request international merchandise trade statistics from countries only in one commodity classification — the most recent one in use by a country. Any recompilation needed by an international organization is done by that organization. This means that countries are not required to maintain concordances and recompile international merchandise trade statistics according to various versions of HS and SITC for international reporting;

(b) The problem of maintaining comparable long-term series is addressed in paragraphs 5, 6 and 7 above; the basic point is that such time series can normally be maintained only in terms of the oldest, least detailed commodity classification i.e., SITC, Rev. 1. Data in terms of SITC, Rev. 1 cannot be well converted in SITC, Rev. 2, particularly at the 4- and 5-digit levels; and even at the 3-digit level, many items do not convert well. Even conversion from the newer to the older classification cannot always be well done if the principles of classification change between versions of the same classification. For example, although SITC, Rev. 2 can be converted to SITC, Rev. 1 quite well, SITC, Rev. 3 conversion to SITC, Rev. 2 is problematic for some codes, and data so converted are not fully comparable with data for earlier years reported in terms of SITC, Rev. 2, partly because SITC, Rev. 3 was modelled closely on HS 88, which was adopted in the same year, and used different classification criteria than did SITC, Rev. 2. These kinds of problems are not unique to international merchandise trade statistics. Part of the solution in general is to adopt as a criterion, when revising a classification, the need to take account of continuity in time series;

(c) As concerns the use of CPC (goods part) as a classification for international merchandise trade statistics, the Commission will recall that the current explicit recommendation of the Commission is that countries should compile and disseminate their international merchandise trade statistics according to HS, and that the Commission decided that it would consider the use of CPC (goods part) in place of SITC, Rev. 3 for analytical purposes when the CPC had been revised and evaluated.[2] The experience of the United Nations Statistics Division experience to date is that (i) there have been no requests from users for international merchandise trade statistics according to CPC; (ii) no national statistical offices are known to be publishing their international merchandise trade statistics according to CPC; however, (iii) some international organizations have indicated that they would find trade data according to CPC useful. Since HS data exists in Comtrade and the Division has developed a conversion table for HS to CPC, a stand-alone database of trade data according to CPC (goods part) could be developed on diskette for the purposes of evaluating that part of the classification itself and its use (for example, 20 countries for each of 1997 and 1990). In the longer term, depending upon the outcome of the evaluation and revision of CPC and on user needs, provision could be made for data according to CPC (goods part) to be added to Comtrade, as set out in paragraph 14 above (for further details on CPC and its revision and evaluation in the overall context of economic and social classifications, see E/CN.3/1999/17 and E/CN.3/1999/18).

V. Points for discussion

17. Since there are users of the international merchandise trade statistics according to each of the five commodity

classifications in Comtrade, and since disk space usage and its cost and the amount of data processing have been reduced, so that the need to reduce the number of classifications in the Database is not acute, the United Nations Statistics Division proposes that:

(a) The Commodity Trade Statistics Database be maintained in its current form and content for the immediate future;

(b) In the short term, the Division produce a data set of merchandise trade according to CPC (goods part) on diskette for reviewing that classification and its use;

(c) In the longer term, the Division investigate the ways in which trade statistics according to CPC (and ISIC) could be made available if a sufficient need for that data were to be established.

Notes

[1] *Official Records of the Economic and Social Council, 1995, Supplement No. 8* (E/1995/28), para. 19 (f).

[2] *Official Records of the Economic and Social Council, 1993, Supplement No. 6* (E/1993/26), paras. 162 (d) and (g).